

# FREE EXPONENTIAL FAMILIES AS KERNEL FAMILIES

WŁODZIMIERZ BRYC

**ABSTRACT.** Free exponential families have been previously introduced as a special case of the  $q$ -exponential family. We show that free exponential families arise also from the approach analogous to the definition of exponential families by using the Cauchy-Stieltjes kernel  $1/(1-\theta x)$  instead of the exponential kernel  $\exp(\theta x)$ . We use this approach to re-derive some known results and to study further similarities with exponential families and reproductive exponential models.

## 1. INTRODUCTION

Since the seminal work of Voiculescu [23], there has been a flurry of activity on how properties of free convolution  $\mu \boxplus \nu$  of probability measures are similar to and how they differ from properties of classical convolution  $\mu * \nu$ . In particular, free probability analogues of the Central Limit Theorem, of the Poisson limit theorem, and the Lévy-Khinchin representation of  $\boxplus$ -infinitely divisible laws are now known, see [13]. New additional analogies between free and classical probability are developed in [4, 5]. In this paper we study a free probability analogue of the concept of exponential family.

Free exponential families were introduced in [10, Definition 4.1] as part of a study of the relations between approximation operators, classical exponential families and their  $q$ -deformations. An alternative approach to free exponential families which we adopt in this paper emphasizes similarities to classical exponential families, and is based on an idea of kernel family introduced in [25]. We show that the two approaches are closely related, and that every non-degenerate compactly supported probability measure generates a free exponential family, see Theorem 3.1. We then relate variance functions of free exponential

---

*Date:* September 26, 2008. Revised: April 24, 2009.

*2000 Mathematics Subject Classification.* 60E10; 46L54.

*Key words and phrases.* exponential families, dispersion models, free convolution, free cumulants.

Research partially supported by NSF grant #DMS-05-04198 and by Taft Research Seminar 2008/09.

families to free cumulants. This relation is simpler than the corresponding relation for classical exponential families and is expressed by a concise formula. We apply the formula to compute free cumulants of the "free gamma" law which were stated without proof in [9], to derive simple necessary conditions for a smooth function to be the variance function of a free exponential family, and to investigate similarities with classical dispersion models [15].

## 2. CAUCHY-STIELTJES KERNEL FAMILIES

According to Wesolowski [25], the kernel family generated by a kernel  $k(x, \theta)$  consists of the probability measures

$$\{k(x, \theta)/L(\theta)\nu(dx) : \theta \in \Theta\},$$

where  $L(\theta) = \int k(x, \theta)\nu(dx)$  is the normalizing constant, and  $\nu$  is the generating measure.

The theory of exponential families is based on the kernel  $k(x, \theta) = e^{\theta x}$ . See, e.g., [15], [16], or [12, Section 2.3]. In this paper, we consider the Cauchy-Stieltjes kernel

$$k(x, \theta) = \frac{1}{1 - \theta x}.$$

**Definition 2.1.** Suppose  $\nu$  is a compactly supported non-degenerate (i.e. not a point mass) probability measure. Let

$$M(\theta) = \int \frac{1}{1 - \theta x} \nu(dx).$$

The Cauchy-Stieltjes family generated by  $\nu$  is the family of probability measures

$$\mathcal{K}(\nu; \Theta) = \left\{ P_\theta(dx) = \frac{1}{M(\theta)(1 - \theta x)} \nu(dx) : \theta \in \Theta \right\}, \quad (2.1)$$

where  $\Theta \ni 0$  is an open set on which  $M(\theta)$  is well defined, strictly positive and  $\theta \text{ supp}(\nu) \subset (-\infty, 1)$ . (We shall only consider  $\Theta = (-\varepsilon, \varepsilon)$  with  $\varepsilon > 0$  small enough.)

Our first goal is to show that the Cauchy-Stieltjes family is essentially the same concept as the concept of free exponential family introduced in [10]. We begin with a suitable reparametrization of  $\mathcal{K}(\nu; \Theta)$ .

**2.1. Parameterizations by the mean.** From (2.1) we compute the mean  $m(\theta) = \int x P_\theta(dx)$ . Since  $P_0 = \nu$  we get  $m(0) = \int x \nu(dx) = m_0$  and for  $\theta \neq 0$ , a calculation gives

$$m(\theta) = \frac{M(\theta) - 1}{\theta M(\theta)}. \quad (2.2)$$

Since  $M(0) = 1$  and  $M(\theta)$  is analytic at  $\theta = 0$ , we see that  $m(\theta)$  is analytic for  $|\theta|$  small enough. We have

$$m'(\theta) = \frac{M(\theta) + \theta M'(\theta) - (M(\theta))^2}{\theta^2 (M(\theta))^2}.$$

Since  $\nu$  is non-degenerate,

$$\begin{aligned} & M(\theta) + \theta M'(\theta) - (M(\theta))^2 \\ &= \int \frac{1}{(1 - \theta x)^2} \nu(dx) - \left( \int \frac{1}{1 - \theta x} \nu(dx) \right)^2 > 0 \end{aligned} \quad (2.3)$$

for all  $|\theta| > 0$  small enough. Thus the function  $\theta \mapsto m(\theta)$  is increasing on an open interval containing 0. Denoting by  $\psi$  the inverse function, we are thus lead to parametrization of a subset of  $\mathcal{K}(\nu, \Theta)$  by the mean,

$$\mathcal{L}(\nu; R) = \{Q_m(dx) = P_{\psi(m)}(dx) : m \in R\}, \quad (2.4)$$

where  $R = m(\Theta_0)$ , and  $\Theta_0 \subset \Theta$  is an appropriate interval with  $0 \in \Theta_0$ . Notice that we refrain from claiming that (2.1) and (2.4) are equivalent: we only claim that for any pair of open sets  $\Theta \ni 0$  and  $R \ni m_0$ , there are open sets  $\Theta_0 \ni 0$  and  $R_0 \ni m_0$  such that  $\mathcal{L}(\nu; R_0) \subset \mathcal{K}(\nu; \Theta)$ , and  $\mathcal{K}(\nu; \Theta_0) \subset \mathcal{L}(\nu; R)$ .

The variance function of the Cauchy-Stieltjes family (2.4) is

$$V(m) = \int (x - m)^2 Q_m(dx). \quad (2.5)$$

### 3. RELATION TO FREE EXPONENTIAL FAMILIES

The following generalizes slightly [10, Section 4]; note that this definition is not constructive: for a given  $V$ , the corresponding free exponential family may fail to exist, see Example 3.2.

**Definition 3.1.** The free exponential family with variance function  $V$  generated by a compactly supported measure  $\nu$  with mean  $m_0 \in (A, B)$  is a family of probability measures

$$\mathcal{F}_{m_0}(V) = \left\{ \frac{V(m)}{V(m) + (m - m_0)(m - x)} \nu(dx) : m \in (A, B) \right\}. \quad (3.1)$$

The next result shows that Cauchy-Stieltjes kernel families under parametrization by the mean are essentially the same as free exponential families, thus providing existence argument for free exponential families. Furthermore, the generating measure  $\nu$  is determined uniquely by  $m_0$  and the variance function  $V(m)$ ; the latter is an analog of the classical uniqueness theorem for exponential families, see [15, Theorem 2.11] or [17, Proposition 2.2].

Recall that the Cauchy-Stieltjes transform of a probability measure  $\nu$  is

$$G_\nu(z) = \int \frac{1}{z-x} \nu(dx). \quad (3.2)$$

If  $\nu$  is compactly supported then  $G_\nu$  is analytic in the neighborhood of  $\infty$  in the complex plane; in particular, compactly supported measures are determined uniquely by  $G_\nu(z)$  for large enough real  $z$ .

**Theorem 3.1.** *Let  $\{P_\theta : \theta \in \Theta\}$  be a Cauchy-Stieltjes family (2.1) generated by a non-degenerate compactly supported measure  $\nu$  with  $\int x\nu(dx) = m_0$ . Then there is a neighborhood  $(A, B)$  of  $m_0$  in which the variance function  $V$  in (2.5) is analytic, strictly positive, and*

$$P_{\psi(m)}(dx) = \frac{V(m)}{V(m) + (m - m_0)(m - x)} \nu(dx), \quad m \in (A, B). \quad (3.3)$$

*Conversely, if  $V$  is analytic and strictly positive in a neighborhood of  $m_0$ , and there is a probability measure  $\nu$  with mean  $m_0$  such that the (positive) measures*

$$Q_m(dx) = \frac{V(m)}{V(m) + (m - m_0)(m - x)} \nu(dx) \quad (3.4)$$

*are probability measures for all  $m$  in a neighborhood of  $m_0$ , then  $\nu$  is compactly supported, non-degenerate, and is determined uniquely by (3.2) with*

$$G_\nu(z) = \frac{m - m_0}{V(m)}, \quad (3.5)$$

*with  $z = 1/\psi(m) = m + V(m)/(m - m_0)$ . In particular, in a perhaps smaller neighborhood  $R$  of  $m_0$ , probability measures (3.4) for  $m \in R$  are included in the Cauchy-Stieltjes family (2.4) generated by  $\nu$ .*

*Proof.* We first calculate the variance  $v(\theta) = \int x^2 P_\theta(dx) - m^2(\theta) = \int x(x - m_0) P_\theta(dx) - m(\theta)(m(\theta) - m_0)$ . Since

$$\begin{aligned} \int x(x - m_0) P_\theta(dx) &= \int \frac{(x - m_0)(\theta x - 1) + x - m_0}{\theta M(\theta)(1 - \theta x)} \nu(dx) \\ &= \frac{m(\theta) - m_0}{\theta}, \end{aligned}$$

we see that for  $\theta \neq 0$  the variance is

$$v(\theta) = (m(\theta) - m_0) \left( \frac{1}{\theta} - m(\theta) \right). \quad (3.6)$$

Since  $m(\theta)$  is analytic at  $\theta = 0$  and  $m(0) = m_0$ , this shows that  $v(\theta)$  is analytic at  $\theta = 0$ . Let  $V(m) = v(\psi(m))$  denote the variance

function in parametrization of (a subset of)  $\mathcal{K}$  by the mean; clearly  $V$  is an analytic function in a neighborhood of  $m = m_0$ .

Now (3.6) implies  $(m(\theta) - m_0)/v(\theta) = \theta/(1 - \theta m(\theta))$  and (2.2) implies  $1 - \theta m(\theta) = 1/M(\theta)$ , so (3.4) is equivalent to

$$\begin{aligned} Q_{m(\theta)}(dx) &= \frac{1}{1 + (m(\theta) - x)(m(\theta) - m_0)/v(\theta)} \nu(dx) \\ &= \frac{1 - \theta m(\theta)}{1 - \theta m(\theta) + (m(\theta) - x)\theta} \nu(dx) = \frac{1 - \theta m(\theta)}{1 - \theta x} \nu(dx) \\ &= \frac{1}{M(\theta)(1 - \theta x)} \nu(dx) = P_\theta(dx). \end{aligned}$$

To prove the converse implication, note that for  $m$  such that  $V(m) > 0$  we can re-write  $\int Q_m(dx) = 1$  as

$$\int \frac{1}{V(m)/(m - m_0) + m - x} \nu(dx) = \frac{m - m_0}{V(m)}.$$

Thus with

$$z = m + \frac{V(m)}{m - m_0}, \quad (3.7)$$

we get (3.5). Since  $\lim_{m \rightarrow m_0^\pm} V(m)/(m - m_0) = \pm\infty$ , this shows that Cauchy-Stieltjes transform  $G_\nu(z)$  is defined for all real  $z$  with  $|z|$  large enough. This implies that  $\nu$  has compact support, with moments that are uniquely determined from the corresponding moment generating function  $M(z) = 1/zG(1/z)$  for  $z$  small enough. (Compactness of support is also proved more directly in the proof of Theorem 3.3.) Finally,  $\nu$  is non-degenerate as its variance is  $V(m_0) > 0$ .  $\square$

**Remark 3.1.** Solving equation (3.6) for  $\theta$  we see that

$$\psi(m) = \frac{m - m_0}{m(m - m_0) + V(m)}.$$

Thus a necessary condition for  $V$  to be a variance function is that  $m \mapsto m + V(m)/(m - m_0)$  is decreasing in a neighborhood of  $m_0$ , see (2.3).

**3.1. Free exponential families with quadratic variance function.** In this section we recall [10, Theorem 4.2]; since manuscript [10] is available in preprint form only and we have already set up all identities needed for the proof, we include the argument which is taken from [10]. The corresponding result for classical exponential families is [14, Theorem 3.3] and [18, Section 4]; the result for  $q$ -exponential families is [10, Theorem 3.2].

**Theorem 3.2** ([10, Theorem 4.2]). *Suppose  $b \geq -1$ ,  $m_0 = 0$ . The free exponential family with the variance function*

$$V(m) = 1 + am + bm^2$$

*consists of probability measures (3.4) with generating measure*

$$\nu(dx) = \frac{\sqrt{4(1+b) - (x-a)^2}}{2\pi(bx^2 + ax + 1)} 1_{(a-2\sqrt{1+b}, a+2\sqrt{1+b})} dx + p_1 \delta_{x_1} + p_2 \delta_{x_2}, \quad (3.8)$$

*where the discrete part of  $\nu$  is absent except for the following cases:*

- (i) *if  $b = 0$ ,  $a^2 > 1$ , then  $p_1 = 1 - 1/a^2$ ,  $x_1 = -1/a$ ,  $p_2 = 0$ .*
- (ii) *if  $b > 0$  and  $a^2 > 4b$ , then  $p_1 = \max\left\{0, 1 - \frac{|a| - \sqrt{a^2 - 4b}}{2b\sqrt{a^2 - 4b}}\right\}$ ,  $p_2 = 0$ , and  $x_1 = \pm \frac{|a| - \sqrt{a^2 - 4b}}{2b}$  with the sign opposite to the sign of  $a$ .*
- (iii) *if  $-1 \leq b < 0$  then there are two atoms at*

$$x_{1,2} = \frac{-a \pm \sqrt{a^2 - 4b}}{2b}, \quad p_{1,2} = 1 + \frac{\sqrt{a^2 - 4b} \mp a}{2b\sqrt{a^2 - 4b}}.$$

*Proof.* With  $m_0 = 0$  and  $V(m) = 1 + am + bm^2$ , equation (3.7) can be solved for  $m$ , giving

$$m = \frac{z - a - \sqrt{(a - z)^2 - 4(1 + b)}}{2(1 + b)},$$

so (3.5) gives

$$G_\nu(z) = \frac{a + z + 2bz - \sqrt{(a - z)^2 - 4(1 + b)}}{2(1 + az + bz^2)}. \quad (3.9)$$

This Cauchy-Stieltjes transform corresponds to the free-Meixner law (3.8), see [1, 21].  $\square$

Theorem 3.2 results covers a number of important laws that appeared in the literature. Up to a dilation and convolution with a degenerate law  $\delta_a$  (i.e. up to "the type") the generating measure  $\nu$  is:

- (i) the Wigner's semicircle (free Gaussian) law if  $a = b = 0$ ; see [24, Section 2.5];
- (ii) the Marchenko-Pastur (free Poisson) type law if  $b = 0$  and  $a \neq 0$ ; see [24, Section 2.7];
- (iii) the free Pascal (free negative binomial) type law if  $b > 0$  and  $a^2 > 4b$ ; see [21, Example 3.6];
- (iv) the free Gamma type law if  $b > 0$  and  $a^2 = 4b$ ; see [9, Proposition 3.6];

- (v) the free analog of hyperbolic type law if  $b > 0$  and  $a^2 < 4b$ ; see [1, Theorem 4];
- (vi) the free binomial type law if  $-1 \leq b < 0$ ; see [21, Example 3.4] and [9, Proposition 2.1].

The laws in (i)-(v) are infinitely divisible with respect to free additive convolution (we recall the definition near (3.17)). In [1, Theorem 4] they appear in connection to martingale polynomials with respect to free Lévy processes; free infinite divisibility is analyzed also in [21]; [2] studies further free probability aspects of this family; in [9, Theorem 3.2] the same laws appear as a solution to a quadratic regression problem in free probability; in [11, Theorem 4.3] these laws occur in a “classical regression” problem.

**3.2. Free Cumulants and Variance Functions.** Recall that if  $\nu$  is a compactly supported measure with the Cauchy-Stieltjes transform  $G_\nu$ , then the inverse function  $K_\nu(z) = G_\nu^{-1}(z)$  exists for small enough  $z \neq 0$ , see [24]. The  $R$ -transform is defined as

$$R_\nu(z) = K_\nu(z) - 1/z \quad (3.10)$$

and is analytic at  $z = 0$ ,

$$R_\nu(z) = \sum_{n=1}^{\infty} c_n(\nu) z^{n-1}. \quad (3.11)$$

The coefficients  $c_n = c_n(\nu)$  are called free cumulants of measure  $\nu$ , see [22].

The following result extends [10, Remark 4.4] and plays a role analogous to Jørgensen’s theorem [17, Theorem 3.2]. (For the formula connecting classical cumulants with the variance functions of natural exponential families, see [18, (2.10)] or [15, Exercise 2.14].)

**Theorem 3.3.** *Suppose  $V$  is analytic in a neighborhood of  $m_0$ ,  $V(m_0) > 0$ , and  $\nu$  is a probability measure with finite all moments, such that  $\int x\nu(dx) = m_0$ . Then the following conditions are equivalent.*

- (i)  $\nu$  is non-degenerate, compactly supported, and there exists an interval  $(A, B) \ni m_0$  such that (3.1) defines a family of probability measures parameterized by the mean with the variance function  $V$ .
- (ii) The free cumulants (3.11) of  $\nu$  are  $c_1 = m_0$ , and for  $n \geq 1$

$$c_{n+1} = \frac{1}{n!} \frac{d^{n-1}}{dx^{n-1}} (V(x))^n \Big|_{x=m_0}. \quad (3.12)$$

*Proof.* Suppose that  $V$  determines the free exponential family generated by a compactly supported measure  $\nu$ . For  $m \neq 0$  close enough to 0, from (3.5) and (3.7) we get

$$m + \frac{V(m)}{m - m_0} = K_\nu \left( \frac{m - m_0}{V(m)} \right).$$

Thus, (3.10) says that the  $R$ -transform of  $\nu$  satisfies

$$R_\nu \left( \frac{m - m_0}{V(m)} \right) = m. \quad (3.13)$$

From this we derive (3.12) by using the Lagrange expansion theorem, which says that if  $\phi(z)$  is analytic in a neighborhood of  $z = m_0$ ,  $\phi(m_0) \neq 0$  and  $\xi := (m - m_0)/\phi(m)$  then

$$m(\xi) = m_0 + \sum_{n=1}^{\infty} \frac{\xi^n}{n!} \left[ \frac{d^{n-1}[\phi(x)]^n}{dx^{n-1}} \right]_{x=m_0}. \quad (3.14)$$

(See, e.g., [20, (L), page 145].)

Suppose now that a probability measure  $\nu$  satisfies (3.12) and  $\int x\nu(dx) = m_0$ . Then the variance  $c_2(\nu) = V(m_0) > 0$ , so  $\nu$  is non-degenerate. We first verify that  $\nu$  has compact support. Since  $V$  is analytic, (3.12) implies that

$$c_{n+1} = \frac{1}{2\pi ni} \oint_{|z-m_0|=\delta} V(z)^n / (z - m_0)^{n-1}, \quad (3.15)$$

so there exist  $M > 0$  such that  $|c_n| \leq M^n$  for all  $n \geq 1$ . The compactness of support follows now from [6, Corollary 1.6]; for completeness we include the proof. Denoting by  $\mathcal{NC}[n]$  the set of non-crossing partitions of  $\{1, 2, \dots, n\}$ , from [13, (2.5.8)] we have

$$\int x^{2n} \nu(dx) = \sum_{\nu \in \mathcal{NC}[2n]} \prod_{B \in \nu} c_{|B|} \leq M^{2n} \# \mathcal{NC}[2n] = M^{2n} \frac{1}{2n+1} \binom{4n}{2n};$$

for the last equality, see [13, (2.5.11)]. Since the  $m$ -th Catalan number is less than  $4^m$ ,

$$\limsup_{n \rightarrow \infty} \left( \int |x|^{2n} \nu(dx) \right)^{1/(2n)} \leq 4M < \infty,$$

and  $\nu$  has compact support.

From  $\text{supp}(\nu) \subset [-4M, 4M]$  we deduce that the Cauchy-Stieltjes transform  $G_\nu(z)$  is analytic for  $|z| > 4M$ , and the  $R$ -series is analytic for all  $|z|$  small enough.



Since  $V(m) > 0$  for  $m$  close enough to  $m_0$ , taking the derivative we see that  $z \mapsto (z - m_0)/V(z)$  is increasing in a neighborhood of  $z = m_0$ . Denoting by  $h$  the inverse, we have

$$h\left(\frac{z - m_0}{V(z)}\right) = z.$$

From  $c_1(\nu) = m_0$  we see that  $R(0) = m_0 = h(0)$ . By (3.14), we see that all derivatives of  $h$  at  $z = 0$  match the derivatives of  $R$ . Thus  $h(z) = R(z)$  and (3.13) holds for all  $m$  in a neighborhood of 0. For analytic  $G_\nu$ , the latter is equivalent to (3.4) holding for all  $m$  close enough to 0. Thus  $V(m)$  is the variance function of a free exponential family generated by  $\nu$  with  $m \in (-\delta, \delta)$  for some  $\delta > 0$ .  $\square$

We now use (3.12) to relate certain free cumulants to Catalan numbers.

**Corollary 3.4.** [9, Remark 5.7] *If  $\nu$  is the standardized free gamma Meixner law, i.e. it generates the free exponential family with  $m_0 = 0$  and variance function  $V(m) = (1 + am)^2$ , then its free cumulants are*

$$c_{k+1}(\nu) = \frac{1}{k+1} \binom{2k}{k} a^{k-1}, \quad k \geq 1.$$

This fact was stated without proof in [9, Remark 5.7]; the approach indicated there lead to a relatively long proof.

*Proof.* From (3.12),

$$\begin{aligned} c_{k+1}(\nu) &= \frac{1}{k!} \frac{d^{k-1}}{dx^{k-1}} (1 + ax)^{2k} \Big|_{x=0} = a^{k-1} \frac{2k(2k-1) \dots (k+2)}{k!} \\ &= \frac{a^{k-1}}{k+1} \frac{(2k)!}{(k!)^2}. \end{aligned}$$

$\square$

Recall that the free additive convolution of compactly supported probability measures  $\mu, \nu$  is a unique compactly supported measure denoted by  $\mu \boxplus \nu$  with the  $R$ -transform

$$R_{\mu \boxplus \nu}(z) = R_\mu(z) + R_\nu(z). \quad (3.16)$$

(See [23].) Equivalently, free cumulants linearize free convolution,

$$c_n(\mu \boxplus \nu) = c_n(\mu) + c_n(\nu), \quad n \geq 1 \quad (3.17)$$

just like classical cumulants linearize the classical convolution. Recall that  $\mu$  is  $\boxplus$ -infinitely divisible if for every  $n = 1, 2, \dots$ , there is a measure  $\nu$  such that  $\mu = \nu \boxplus \nu \boxplus \dots \boxplus \nu$  (the  $n$ -fold free convolution).

**Corollary 3.5.**  $V(m) = 1/(1-m)$  is a variance function of a free exponential family generated by the centered  $\boxplus$ -infinitely divisible measure  $\nu$  with free cumulants

$$c_{k+1}(\nu) = \frac{1}{k} \binom{2k-2}{k-1}, \quad k \geq 1.$$

*Proof.* From (3.12),

$$c_{k+1} = \frac{1}{k!} \frac{d^{k-1}}{dx^{k-1}} (1-x)^{-k} \Big|_{x=0} = \frac{k(k+1) \dots (2k-2)}{k!} = \frac{1}{k} \frac{(2k-2)!}{((k-1)!)^2}.$$

It is well known that Catalan numbers are even moments of the semi-circle law,

$$c_{k+1} = \int_{-2}^2 x^{2k} \frac{\sqrt{4-x^2}}{2\pi} dx = \int_0^4 x^k \frac{\sqrt{4/x-1}}{2\pi} dx.$$

Therefore,  $R(z) = \sum_{k=1}^{\infty} c_k z^{k-1} = \int_0^4 \frac{zx}{1-zx} \frac{\sqrt{4/x-1}}{2\pi} dx = \int_0^4 \frac{z}{1-zx} \frac{\sqrt{4x-x^2}}{2\pi} dx$  corresponds to  $\boxplus$ -infinitely divisible law, see [13, Theorem 3.3.6]. Thus Catalan numbers  $c_{k+1}$  with  $c_1 = 0$  are indeed free cumulants of some  $\boxplus$ -infinitely divisible measure  $\nu$ .  $\square$

It is known that not every function  $V$  is a variance function of a natural exponential family. It is therefore not surprising that not every analytic functions  $V$  can serve as the variance functions for a free exponential family.

**Corollary 3.6.** Suppose  $V$  is analytic at 0 and  $V(0) = 1$ ,  $V''(0) < -2$ . Then  $V$  cannot be a variance function of a free exponential family with  $m_0 = 0$ .

*Proof.* Suppose  $V$  generates a free exponential family with generating measure  $\nu$ . Let  $m_j = \int x^j \nu(dx)$  with  $m_1 = 0$ ,  $m_2 = 1$ . Then the  $3 \times 3$  Hankel determinant is

$$\det \begin{bmatrix} 1 & m_1 & m_2 \\ m_1 & m_2 & m_3 \\ m_2 & m_3 & m_4 \end{bmatrix} = \det \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & m_3 \\ 1 & m_3 & m_4 \end{bmatrix} = m_4 - m_3^2 - 1 \geq 0.$$

Using (3.12), the fourth moment is

$$m_4 = c_4(\nu) + 2c_2^2(\nu) = c_4(\nu) + 2 = V'(0)^2 + \frac{1}{2}V''(0) + 2$$

and  $m_3 = c_3(\nu) = V'(0)$ , see [13, (2.5.8)]. Thus  $m_4 - m_3^2 - 1 \geq 0$  translates into  $V''(0) \geq -2$ .  $\square$

**Example 3.1.** If  $b < -1$ , then  $V(m) = 1 + am + bm^2$  is not a variance function of a free exponential family with  $m_0 = 0$ . Compare Theorem 3.2. (This can also be seen from Remark 3.1.)

**Example 3.2.**  $V(m) = (1 - m)/(1 + m)$  is not a variance function of a free exponential family with  $m_0 = 0$ . (This can also be seen from Remark 3.1.)

Combining (3.12) with the  $\boxplus$ -Lévy-Khinchin formula [13, Theorem 3.3.6], compare [3, Lemma 3.4], we get also the following.

**Corollary 3.7.** *Suppose  $V(m)$  is analytic at 0,  $V(0) = 1$ . Then the following conditions are equivalent.*

- (i) *There exists a centered  $\boxplus$ -infinitely divisible probability measure  $\nu$  such that  $V$  is the variance function of a free exponential family generated by  $\nu$ .*
- (ii) *There exists a compactly supported probability measure  $\omega$  such that*

$$\frac{1}{n!} \frac{d^{n-1}}{dx^{n-1}} (V(x))^n \Big|_{x=0} = \int x^{n-1} \omega(dx), \quad n \geq 1. \quad (3.18)$$

The Cauchy-Schwarz inequality applied to the right hand side of the Lévy-Khinchin formula (3.18) implies  $(V^3)''/6 \geq ((V^2)')^2/4$ . This gives a simple necessary condition.

**Corollary 3.8.** *If  $V$  is analytic at 0,  $V(0) = 1$ ,  $V''(0) < 0$  then  $V$  cannot be the variance function of a free exponential family generated by a centered  $\boxplus$ -infinitely divisible measure.*

We remark that the bound is sharp: from Theorem 3.2 we see that  $V(m) = 1$  is a variance function of the free exponential family generated by the semicircle law; all of its members are infinitely divisible, see Example 4.1.

**Example 3.3** (Compare [21, Theorem 3.2]). If  $b < 0$  then  $V(m) = 1 + am + bm^2$  cannot be the variance function of a free exponential family generated by a centered  $\boxplus$ -infinitely divisible measure.

**3.3. Reproductive property.** Natural exponential families have two "reproductive" properties. The first one is usually not named, and says that if a compactly supported measure  $\nu$  generates natural exponential family  $\mathcal{F}$  and  $\mu \in \mathcal{F}(\nu)$  then  $\mathcal{F}(\mu) = \mathcal{F}$ . This is usually interpreted as a statement that the natural exponential family  $\mathcal{F}$  is determined solely by the variance function  $V$  and can have many generating measures.

The analog of this property fails for free exponential families due to the fact that the generating measure is determined uniquely by the variance function and parameter  $m_0$ . For example, a free exponential family  $\mathcal{F}$  generated by the centered semicircle law consists of the affine transformations of the Marchenko-Pastur laws, and for  $m_0 \neq 0$  the free exponential family generated by  $\mu \in \mathcal{F}$  with mean  $m_0$  contains no other measures in common with  $\mathcal{F}$  except for  $\mu$ .

The second property which in [15, (3.16)] is indeed called the reproductive property of an exponential family states that if  $\mu \in \mathcal{F}(V)$ , then for all  $n \in \mathbb{N}$  the law of the sample mean,  $D_n(\mu^{*n})$ , is in  $\mathcal{F}(V/n)$ . Here  $D_r(\mu)(U) := \mu(rU)$  denotes the dilation of measure  $\mu$  by a number  $r \neq 0$ ; in probabilistic language, if  $\mathcal{L}(X) = \mu$  then  $\mathcal{L}(X/r) = D_r(\mu)$ .

Our goal is to prove an analogue of this result for the Cauchy-Stieltjes families.

Let  $\mu^{\boxplus r}$  denote the  $r$ -fold free additive convolution of  $\mu$  with itself. In contrast to classical convolution, this operation is well defined for all real  $r \geq 1$ , see [19].

**Proposition 3.9** ([10, Proposition 4.3]). *If a function  $V$  analytic at  $m_0$  is a variance function of a free exponential family generated by a compactly supported probability measure  $\nu$  with  $m_0 = \int x\nu(dx)$ , then for each  $\lambda \geq 1$  there exists a neighborhood of  $m_0$  such that  $V/\lambda$  is the variance function of the free exponential family generated by measure*

$$\nu_\lambda := D_\lambda(\nu^{\boxplus \lambda}).$$

*Moreover, if for each  $\lambda > 0$ , there is a neighborhood of  $m_0$  such that  $V/\lambda$  is a variance function of some free exponential family, then  $\nu$  is  $\boxplus$ -infinitely divisible.*

We note that in contrast to classical natural exponential families, the neighborhood of  $m_0$  where  $m \mapsto V(m)/\lambda$  is a variance function may vary with  $\lambda$ , see Example 4.1.

*Proof.* Combining (3.12) with  $R_{aX+b}(z) = b + aR_X(az)$ , we see that the free cumulants of  $\nu_\lambda$  are  $c_1(\nu_\lambda) = c_1(\nu) = m_0$  and for  $n \geq 1$

$$c_{n+1}(\nu_\lambda) = \frac{1}{\lambda^n} c_{n+1}(\nu) = \frac{1}{n!} \frac{d^{n-1}}{dx^{n-1}} \left( \frac{V(x)}{\lambda} \right)^n \Big|_{x=m_0}.$$

Theorem 3.3 implies that  $V/\lambda$  is the variance function of the free exponential family generated by  $\nu_\lambda$ .

If  $\nu_{1/n}$  exists for all  $n \in \mathbb{N}$ , then the first part of the proposition together with uniqueness theorem (Theorem 3.1) implies that  $\nu = (D_n(\nu_{1/n}))^{\boxplus n}$ , proving  $\boxplus$ -infinite divisibility.  $\square$

## 4. MARCHENKO-PASTUR APPROXIMATION

Let

$$\omega_{a,\sigma}(dx) = \frac{\sqrt{4\sigma^2 - (x-a)^2}}{2\pi\sigma^2} 1_{|x-a| < 2\sigma} dx$$

denote the semicircle law of mean  $a$  and variance  $\sigma^2$ . Up to affine transformations, this is the free Meixner law which appears in Theorem 3.2 as the law which generates the free exponential family  $\mathcal{F}_a(V)$  with the variance function  $V \equiv \sigma^2$ .

Following the analogy with natural exponential families, family  $\mathcal{F}_0(\sigma^2)$  can be thought as a free exponential analog of the normal family. Somewhat surprisingly, this family does not contain all semicircle laws, but instead it contains affine transformations of the (absolutely continuous) Marchenko-Pastur laws.

**Example 4.1** (Semi-circle free exponential family). For  $\lambda > 0$ , let

$$\pi_{m,\lambda}(dx) = \frac{\sqrt{\lambda}\sqrt{4 - \lambda x^2}}{2\pi(1 + \lambda m(m - x))} 1_{x^2 \leq 4/\lambda} dx.$$

Function  $V(m) \equiv 1/\lambda$  is the variance function of the free exponential family

$$\mathcal{F}_0(1/\lambda) = \left\{ \pi_{m,\lambda}(dx) : |m| < 1/\sqrt{\lambda} \right\} \quad (4.1)$$

with the generating measure  $\nu(dx) = \omega_{0,1/\sqrt{\lambda}}(dx)$ .

To verify that the expression integrates to 1 for  $m \neq 0$ , we use the explicit form of the density [13, (3.3.2)] to note that  $\pi_{m,\lambda} = \mathcal{L}(m + 1/(\lambda m) - mX)$  is the law of the affine transformation of a free Poisson (Marchenko-Pastur) random variable  $X$  with parameter  $1/(\lambda m^2)$ . From the properties of Marchenko-Pastur law we see that  $\int \pi_{m,\lambda}(dx) = 1$  iff  $m^2 \leq 1/\lambda$ , so for large  $\lambda$  the interval  $(A, B) \subset (-1/\sqrt{\lambda}, 1/\sqrt{\lambda})$  in (3.1) cannot be chosen independently of  $\lambda$ .

We remark that Biane [7] analyzes  $f \mapsto g(m) := \int f(x)\pi_{m,\lambda}(dx)$  as a mapping of the appropriate Hilbert spaces for complex  $m$ .

We have the following analogue of [15, Theorem 3.4].

**Theorem 4.1** (Marchenko-Pastur approximation). *Suppose the variance function  $V$  of a free exponential family  $\mathcal{F}_{m_0}(V)$  is analytic and strictly positive in a neighborhood of  $m_0$ . Then there is  $\delta > 0$  such that if  $\mathcal{L}(Y_\lambda) \in \mathcal{F}_{m_0}(V/\lambda)$  has mean  $E(Y_\lambda) = m_0 + m/\sqrt{\lambda}$  with  $|m| < \delta$ , then  $\sqrt{\lambda}(Y_\lambda - m_0) \xrightarrow{\mathcal{D}} \pi_{m,1/V(m_0)}$  as  $\lambda \rightarrow \infty$ .*

To prove Theorem 4.1 we will use the following analogue of Mora's Theorem, see [15, Theorem 2.12], or [17, Theorem 2.6].

**Proposition 4.2.** *Suppose  $V_n$  is a family of analytic functions which are variance functions of a sequence of free exponential families  $\{\mathcal{F}_{m_0}(V_n) : n \geq 1\}$ . If  $V_n \rightarrow V$  uniformly in a (complex) neighborhood of  $m_0 \in \mathbb{R}$ , and  $V(m_0) > 0$ , then there is  $\delta > 0$  such that  $V$  is a variance function of a free exponential family  $\mathcal{F}_{m_0}(V)$  parameterized by the mean  $m \in (m_0 - \delta, m_0 + \delta)$ . Moreover, if a sequence of measures  $\mu_n \in \mathcal{F}_{m_0}(V_n)$  is such that  $m_1 = \int x \mu_n(dx) \in (m_0 - \delta, m_0 + \delta)$  does not depend on  $n$ , then  $\mu_n \xrightarrow{\mathcal{D}} \mu$  where  $\mu \in \mathcal{F}_{m_0}(V)$  has the same mean  $\int x \mu(dx) = m_1$ .*

*Proof.* Let  $\nu_n$  be the generating measure for  $\mathcal{F}_{m_0}(V_n)$ . Since  $V_n(z) \rightarrow V(z)$  uniformly in a neighborhood of  $m_0$ , from (3.15) we see that the cumulants  $c_{k+1}(\nu_n)$  converge as  $n \rightarrow \infty$  and  $\sup_n |c_{k+1}(\nu_n)| \leq M^k$  for some  $M < \infty$ . Therefore the  $R$ -transforms of  $\nu_n$  converge to the  $R$ -transform of a compactly supported measure  $\nu$ . Thus  $\nu_n \xrightarrow{\mathcal{D}} \nu$ , and the supports of  $\nu_n$  are uniformly bounded in  $n$ , i.e.,  $\text{supp}(\nu_n) \subset [-A, A]$  for some  $0 < A < \infty$ . By decreasing the value of  $\delta$  we can also ensure that the densities in (3.4) are bounded as functions of  $x \in [-A, A]$  uniformly in  $n$ . So the integrals converge, and  $\nu$  indeed generates a free exponential family with variance  $V$  in a neighborhood of  $m_0$ .

Suppose now  $\mu_n \in \mathcal{F}_{m_0}(V_n)$  and  $\mu \in \mathcal{F}_{m_0}(V)$  have the same mean  $m$  for some  $|m - m_0| < \delta$  small enough. Since the densities in (3.4) are bounded by some constant  $C$  for all  $x \in [-A, A]$ ,  $n \in \mathbb{N}$ , for a bounded continuous function  $g$  we have

$$\begin{aligned} \int g(x) \mu_n(dx) &= \int g(x) \frac{V_n(m)}{V_n(m) + (m - m_0)(m - x)} \nu_n(dx) \\ &= \int g(x) \frac{V(m)}{V(m) + (m - m_0)(m - x)} \nu_n(dx) + \varepsilon_n \\ &\rightarrow \int g(x) \frac{V(m)}{V(m) + (m - m_0)(m - x)} \nu(dx) = \int g(x) \mu(dx) \end{aligned}$$

as  $n \rightarrow \infty$ , where

$$\begin{aligned} |\varepsilon_n| &= \left| \int g(x) \left( \frac{V_n(m)}{V_n(m) + (m - m_0)(m - x)} \right. \right. \\ &\quad \left. \left. - \frac{V(m)}{V(m) + (m - m_0)(m - x)} \right) \nu_n(dx) \right| \\ &\leq C^2 |m - m_0| \sup_{x \in [-A, A]} |g(x)(m - x)| \frac{|V_n(m) - V(m)|}{V_n(m)V(m)} \rightarrow 0. \end{aligned}$$

□

*Proof of Theorem 4.1.* Without loss of generality, we assume  $m_0 = 0$ . Suppose  $\mu \in \mathcal{F}(V)$ . A change of variable shows that  $D_a(\mu)(dx) \in \mathcal{F}(V(am)/a^2)$ , with generating measure  $\nu_a = D_a(\nu)$ . Since

$$\mathcal{L}(Y_\lambda) \in \mathcal{F}\left(\frac{V(m)}{\lambda}\right),$$

this shows that

$$\mathcal{L}(\sqrt{\lambda}Y_\lambda) \in \mathcal{F}\left(V\left(\frac{m}{\sqrt{\lambda}}\right)\right).$$

We now use Proposition 4.2 to the sequence of variance functions  $V_\lambda(m) = V\left(\frac{m}{\sqrt{\lambda}}\right) \rightarrow V(0) = \sigma^2$  as  $\lambda \rightarrow \infty$ . From Proposition 4.2 we deduce that there is  $0 < \delta < \sigma^2$  such that if  $|m| < \delta$  and  $E(\sqrt{\lambda}Y_\lambda) = m$ , then  $\mathcal{L}(\sqrt{\lambda}Y_\lambda) \xrightarrow{\mathcal{D}} \pi_{m,1/\sigma^2} \in \mathcal{F}_0(\sigma^2)$ , see (4.1).  $\square$

By Example 4.1, if  $0 < |m| \leq \sigma$ , then up to affine transformation  $\pi_{m,1/\sigma^2}$  is a Marchenko-Pastur law. Thus in this case Theorem 4.1 gives a Marchenko-Pastur approximation to  $\mathcal{L}(Y_\lambda)$ .

Of course, every compactly supported mean-zero measure  $\nu$  is an element of the Cauchy-Stieltjes family that it generates. Since  $\pi_{0,1/\sigma^2} = \omega_{0,\sigma}$  is the semicircle law, combining Proposition 3.9 with Theorem 4.1 we get the following Free Central Limit Theorem; see [8, 23].

**Corollary 4.3.** *If a probability measure  $\nu$  is compactly supported and centered, then with  $\sigma^2 = \int x^2 \nu(dx)$  we have*

$$D_{\sqrt{n}}(\nu^{\boxplus n}) \xrightarrow{\mathcal{D}} \omega_{0,\sigma}.$$

**Acknowledgements.** The author thanks J. Wesolowski for the copy of [25] and for several helpful discussions. He also thanks the referee for an insightful report which helped to clarify the statement of Theorem 3.1 and led to significant improvement of the paper.

## REFERENCES

- [1] M. Anshelevich. Free martingale polynomials. *J. Funct. Anal.*, 201:228–261, 2003. arXiv:math.CO/0112194.
- [2] M. Anshelevich. Orthogonal polynomials with a resolvent-type generating function. *Trans. Amer. Math. Soc.*, 360(8):4125–4143, 2008. arXiv:math.CO/0410482.
- [3] O. E. Barndorff-Nielsen and S. Thorbjørnsen. Self-decomposability and Lévy processes in free probability. *Bernoulli*, 8(3):323–366, 2002.
- [4] G. Ben Arous and V. Kargin. Free point processes and free extreme values. *Probability Theory and Related Fields*, (to appear), arXiv:0903.2672, 2009.
- [5] G. Ben Arous and D. V. Voiculescu. Free extreme values. *Ann. Probab.*, 34(5):2037–2059, 2006.

- [6] F. Benaych-Georges. Taylor expansions of  $R$ -transforms, application to supports and moments. *Indiana Univ. Math. J.*, 55:465–482, 2006.
- [7] P. Biane. Segal-Bargmann transform, functional calculus on matrix spaces and the theory of semi-circular and circular systems. *J. Funct. Anal.*, 144(1):232–286, 1997.
- [8] M. Bożejko. On  $\Lambda(p)$  sets with minimal constant in discrete noncommutative groups. *Proc. Amer. Math. Soc.*, 51:407–412, 1975.
- [9] M. Bożejko and W. Bryc. On a class of free Lévy laws related to a regression problem. *J. Funct. Anal.*, 236:59–77, 2006. [arxiv.org/abs/math.OA/0410601](https://arxiv.org/abs/math.OA/0410601).
- [10] W. Bryc and M. Ismail. Approximation operators, exponential, and  $q$ -exponential families. Preprint. [arxiv.org/abs/math.ST/0512224](https://arxiv.org/abs/math.ST/0512224), 2005.
- [11] W. Bryc and J. Wesołowski. Conditional moments of  $q$ -Meixner processes. *Probab. Theory Related Fields*, 131:415–441, 2005. [arxiv.org/abs/math.PR/0403016](https://arxiv.org/abs/math.PR/0403016).
- [12] P. Diaconis, K. Khare, and L. Saloff-Coste. Gibbs sampling, exponential families and orthogonal polynomials. *Statistical Science*, 23:151–178, 2008.
- [13] F. Hiai and D. Petz. *The semicircle law, free random variables and entropy*, volume 77 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000.
- [14] M. E. H. Ismail and C. P. May. On a family of approximation operators. *J. Math. Anal. Appl.*, 63(2):446–462, 1978.
- [15] B. Jørgensen. *The theory of dispersion models*, volume 76 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1997.
- [16] G. Letac. *Lectures on natural exponential families and their variance functions*, volume 50 of *Monografías de Matemática [Mathematical Monographs]*. Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 1992.
- [17] G. Letac and M. Mora. Natural real exponential families with cubic variance functions. *Ann. Statist.*, 18(1):1–37, 1990.
- [18] C. N. Morris. Natural exponential families with quadratic variance functions. *Ann. Statist.*, 10(1):65–80, 1982.
- [19] A. Nica and R. Speicher. On the multiplication of free  $N$ -tuples of noncommutative random variables. *Amer. J. Math.*, 118(4):799–837, 1996.
- [20] G. Pólya and G. Szegő. *Problems and theorems in analysis. I*, volume 193 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1978.
- [21] N. Saitoh and H. Yoshida. The infinite divisibility and orthogonal polynomials with a constant recursion formula in free probability theory. *Probab. Math. Statist.*, 21(1):159–170, 2001.
- [22] R. Speicher. Free probability theory and non-crossing partitions. *Sém. Lothar. Combin.*, 39:Art. B39c, 38 pp. (electronic), 1997.
- [23] D. Voiculescu. Addition of certain noncommuting random variables. *J. Funct. Anal.*, 66(3):323–346, 1986.
- [24] D. Voiculescu. Lectures on free probability theory. In *Lectures on probability theory and statistics (Saint-Flour, 1998)*, volume 1738 of *Lecture Notes in Math.*, pages 279–349. Springer, Berlin, 2000.
- [25] J. Wesołowski. Kernel families. Unpublished manuscript, 1999.



DEPARTMENT OF MATHEMATICAL SCIENCES,, UNIVERSITY OF CINCINNATI,,  
PO Box 210025,, CINCINNATI, OH 45221-0025, USA  
*E-mail address:* Wlodzimierz.Bryc@UC.edu